DOCUMENT RESUME

ED 192 345 CS 205 859

AUTHOR Flower, Linda S.: And Others
TITLE Revising Functional Documents: The Scenario Principle. Technical Report No. 10.
INSTITUTION American Institutes for Research, Washington, D.C.: Carnegie-Mellon Univ., Pittsburgh, Pa.; Siegel & Gale, Inc., New York, N.Y.
SPONS AGENCY National Inst. of Education (DHEW), Washington, D.C.
PUB DATE Mar 80
CONTRACT 400-78-0043
NOTE 43p.

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Editing: Government Publications: Guidelines: Language Skills: Research: Technical Writing: *Writing (Composition): Writing Skills
IDENTIFIERS *Document Design Project: *Revision (Written Composition)

ABSTRACT

Defining "functional documents" as those such as regulations, contracts, and policies that people must read in order to act, this paper posits the need for a revision principle by which writers can structure such documents to meet readers' needs. The first part of the paper describes a research project that uncovered some needs of readers in processing and comprehending a functional document. The second part of the paper uses the findings of the research project to formulate a principle for the revision of documents, labeled the "scenario principle," which states that functional prose should provide examples of human agents performing actions in an understandable context. (FL)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED192345

U S DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

DOCUMENT DESIGN PROJECT

CS

Technical Report No. 10

Revising Functional Documents: The Scenario Principle

Linda S. Flower
John R. Hayes
Heidi Swarts

A product of the Document Design Project funded by the National Institute of Education, DHEW, Washington, D.C.

Contract No. NIE-400-78-0043

C 9205 P59

March 1980

CARNEGIE-MELLON UNIVERSITY / 5000 Forbes Avenue, Pittsburgh, PA 15213
with AMERICAN INSTITUTES FOR RESEARCH
and SIEGEL & GALE

Revising Functional Documents: The Scenario Principle

Linda S. Flower, John R. Hayes, and Heidi Swarts

March 1980

This report is part of the work being done under Contract #400-78-0043 of the National Institute of Education. The project officer is Dr. Ramsay Selden of the Teaching and Learning/Reading and Language Group. This report does not, however, necessarily reflect the views of the funding agency.

The project is a combined effort to apply research, teaching, and practical experience to the problem of designing better public documents. We particularly wish to thank Anna-Lena Ericsson and Marshall Atlas for their help in thinking about this problem.

OCT 2 0 1980

ABOUT THE DOCUMENT DESIGN PROJECT

In September 1978, the American Institutes for Research, Carnegie-Mellon University and Siegel & Gale, Inc., began the Document Design Project to foster clear and simple writing and design of public documents. The purpose of the Document Design Project is to help make forms, regulations, brochures, and other written materials easier for people to read, to understand, and to use. Funding for the project comes from the Teaching and Learning/Reading and Language group at the National Institute of Education.

The Project's goal is to increase the knowledge and skills of people who produce public documents. To accomplish this goal, the Document Design Project is

- conducting theoretical and applied research studies on language comprehension, on the ways in which skilled and unskilled writers work, and on problems associated with different document features;
- working directly with government and private agencies as they produce materials for public use;
- developing courses on writing and design for graduate students and undergraduates.

This report is one of a series of technical and explanatory papers resulting from our research.

If you have questions or comments on this report or on other work of the Document Design Project, contact Dr. Janice C. Redish, Director, The Document Design Project, AIR, (202) 342-5071, or Dr. Lee W. Gregg, Department of Psychology, CMU, (412) 578-2781.

TABLE OF CONTENTS

Readability 3

A Reader-Based Structure for Functional Documents 5

Exploring the Needs of Readers 8
A. Method 9
B. Coding System 10
1. Meta Statements 10
2. Content Statements 11
a. Structural statements 11
b. Retrieval statements 12
c. Scenario statements 12
C. Results 13

The Role of Scenarios 16

The Nature of Writers' Revisions 19

Moving Toward A Revision Principle 21

Practical Revision Strategies 26
A. Structuring Information Around the Reader 26
B. Instantiating the Scenario 30

Some Unanswered Questions 32

Conclusion 34

References 36

List of Figures

Figure 1: Categories for Readers' Statements 13
Figure 2: Percentage of Scenarios in Content Statements 15
Figure 3: Frequency of Scenarios in Translate Statements 18
Figure 4: Percentage of Lines with Human Agent/Action Focus 21
Figure 5: Critical Differences Between OLD and REVISED Documents 23

REVISING FUNCTIONAL DOCUMENTS: THE SCENARIO PRINCIPLE

There is one thing on which even politicians can agree: a large number of public documents--from Federal regulations and income tax forms, to insurance policies, loan agreements, and leases--are by and large unreadable, even incomprehensible, to the people who sign and use them. Recently, however, there has been an increased interest in businesses, government, and universities in improving the writing of public documents and in training professional writers or "document designers" how to do so (Charrow, 1979; Redish, 1979; Steinberg, 1979).

As the Clear Language Movement gains momentum, government agencies are rewriting Federal regulations and states are passing laws which require clear language in public documents such as contracts and regulations. But unfortunately, there is little agreement as to what "clear language" or "plain English" really is (Charrow, 1979). As a result, document writers are asking difficult questions, which we could summarize as: "clear" language is a fine goal, but how do you recognize "clear" language when you see it and, given the complexity of our documents and the institutions which produce them, how do you create it? This problem could be stated more operationally as a three-part question: can we define a set of powerful principles which would allow writers to revise public documents so that they meet the needs of the readers who use them? In this paper we will present a set of exploratory experimental studies and

discuss their implications for revising a class of writing we have called "functional documents," that is, documents, such as regulations, contracts, and policies, which people read in order to act. The broad goal to which this study hopes to contribute is, then, the attempt tc offer document writers

- a set of powerful principles
- for making "reader-based" (i.e., designed for the reader)
- revisions

We have defined the problem around this three-part goal for a number of reasons. First, in a large organization such as the government, most writing tasks are essentially revision rather than original composition. For example, regulation writers remake the law into regulations, or rewrite existing regulations, or respond to the changes, additions, and demands of other people in the agency. Therefore, if we could improve the revision prccess, we might have a significant impact on all levels of government writing.

Secondly, it would be easier to revise these documents if writers had a set of principles on which everyone could agree and which were powerful enough to direct the revision process itself. As it is, each document regularly goes through layers of readers and revisers, each of whom makes local revisions based on his or her own independent criteria for "improving" the writing. As a result, these revisions are often contradictory and tend merely

to offer local first aid, such as word changes, even if the document really needs radical surgery.

Finally, we feel that if these revision principles are to be truly effective, they must lead not merely to clear or simple language, but to writing that actively meets the needs of the people who use it, the readers. That is, our goal should be a reader-based revision designed to fit the function of the document and the information processing needs of the readers.

Readability

Currently, the strongest candidate for solving our national "writing problem" appears to be the body of readability formulas, exemplified by the Flesch Reading Ease formula. Readability formulas are a strong contender because they are easy to use, they provide a direct, quantitative measurement of reading speed, and they can predict the reading level required to at least "process" the prose. However, a number of studies have shown that improving readability doesn't reliably affect the reader's comprehension, learning, or retention of the material he or she reads (Klare, 1963, p. 14). Furthermore, readability formulas cannot deal with any units larger than the sentence (Klare, 1963, p. 162). A recent review of readability research reports that it offers "little useful data to writers" on features as important as paragraph construction, organization, and emphasis (Kniffin, 1978, p. 4; see also, Redish, 1980).

Readability formulas are so easy to use because they are so simple. They measure essentially the length of sentences and the familiarity of the vocabulary. As a result, utter nonsense or a paragraph with scrambled sentences could score very high on a readability test, yet be utterly confusing to a real reader (Charrow, 1979). This limitation is critical if, as Wright has said, "The most important characteristic of written information is its structure" (Wright, 1968, p. 448). In support of this emphasis on structure Wright cites a study in which readers' comprehension is improved more dramatically when they are taught how to organize a text as they read than when they are taught either vocabulary or specific retention techniques (Alderman, 1926).

Readability formulas, then, satisfy our first two criteria. They offer an easy to use principle for revising prose. If a passage fails to pass a readability test, one can simply revise it employing more familiar, shorter words and shorter sentences. However, readability, as defined by the formulas, fails to satisfy our third criterion because it fails to meet some of the most important needs of readers. The driving assumption behind readability is that the reader's real problem is processing text and reaching closure. The goal of such revision, then, is to make it easier to hold and relate manageable-sized units in short term memory. Clearly it is very important to write prose which readers can process with relative ease, but it simply isn't

enough. We need a more powerful principle which can organize the information in a document around the larger, comprehension needs of the reader and the function of the document. We need a principle for structural revision.

A Reader-Based Structure for Functional Documents

In considering the structure or organization of a document, we need to consider the whole range of structural options available to the writer, from syntactic structures, which determine the emphasis or focus of a sentence, on up to the hierarchical organization of sentences and of the information itself. In expository prose (e.g., an essay in Time magazine) the key elements in the structure of a paragraph are often its topic sentence and the logical connections between its sentences. In the prose of regulations, however, topic sentences are less significant and the connections between sentences are often procedural rather than strictly logical ones. In such prose, a critical structural feature might be the order of the information: does it move from old to new; or does it lay out a decision path readers could follow to determine how to apply for benefits or to determine if they need to comply with a regulation? Another critical structural feature in such writing is its high-level, overall organization, including the categories into which ideas are grouped and the headings used to define these units. This top-level structure not only helps readers

locate information they need, it sets up a framework for understanding the document as a coherent whole.

Clearly structure is one of the basic ways writers communicate. But how do you define "good" structure in written documents? We are going to argue that for the readers of public documents (1) a good structure is usable (as opposed to other things a structure could be, such as aesthetically pleasing, balanced, or tightly knit) and that (2) a usable structure (or an unusable one) is a key feature in the class of "functional" documents, such as regulations, which we are considering.

From the perspective of cognitive science, the best representation of knowledge on a given topic depends on how that knowledge will be used (Bobrow & Norman, 1975, p. 104). One "use" writers should consider is the way readers will read the document. For example, readers can approach a text with a passive or "bottom-up" reading strategy in which they allow the text to direct their reading process. Or, they can take a more active or "top-down" processing strategy in which their own goals and motives direct how they read and represent the meaning of the text (Wright, 1978; Bobrow & Brown, 1975). It seems likely that readers normally approach public documents with a top-down strategy. They read for a reason: they want to find information, make decisions, and/or act. Consider for example the purpose of the class of documents we have labeled "functional documents," which includes regulations, leases, contracts, loan

agreements, insurance policies, and so on. People rarely read such documents for pleasure or general enlightenment. The purpose of such documents is to regulate action, and people read them in order to act: to comply with a regulatior, apply for benefits, sign or oppose rental terms, purchase or draw on an insurance policy. So from the reader's point of view, good writing is writing which performs its function and offers information structured around that function, not around something else. This relationship between prose structure and the function of the document has been looked at for other kinds of documents. In discussing general expository writing, Flower has argued that one of the key differences between Writer-Based prose and Reader-Based prose is its structure (Flower, 1979). The egocentric focus and narrative or descriptive structure of Writer-Based prose is often quite clear from the writer's point of view, since it is based on the writer's own discovery process or the information before him. But it fails to meet the reader's needs for a hierarchically organized discussion focused on the topic, question, or problem.

In a similar way Kern, Sticht, Welty, and Hauke's excellent work on training manuals demonstrated that the key to effective revision of manuals is also the revision of structure (Kern et al., 1976). The function of a training manual is to tell someone how to do something; therefore a reader-based revision takes that function seriously and organizes information around the task the

reader must perform, such as how to locate land mines, rather than around the general topic of land mines and their locations. A powerful principle for revising manuals, then, is to turn "topic-oriented" prose into procedural or "performance-oriented" prose.

Readability formulas can help, but they can't do the whole job. They need to be complemented by a revision principle which can help writers structure a document around the reader's needs. The first half of this paper will describe a research project intended to uncover some of the things readers need in order to process, or comprehend, a functional document. On the basis of this research we will go on to describe a principle for revision which we have labeled the Scenario principle. In essence, the Scenario principle states that functional prose should be structured around a human agent performing actions in an understandable context. On the face, this may seem like a very common-sensical proposition, yet it would be a radical departure from the current logic and structure of Federal regulations and most contracts.

Exploring the Needs of Readers

Our initial research was guided by two questions:

1. What would a reader-based revision of a Federal regulation look like? What do readers need?
2. Can increased knowledge of what readers need point us

to a practical, powerful principle for revising these regulations and, by implication, other functional documents?

A. Method

Our approach to the first question was straightforward exploration: we collected reading-aloud protocols (tape-recorded transcripts) of subjects reading and paraphrasing the meaning of a Federal regulation as they read. Protocol analysis has proven itself to be a powerful tool for uncovering cognitive processes in writing and other problem solving tasks (Flower & Hayes, 1979) and has recently been used to study reading processes (Waern, 1979). We turned to protocol analysis for this study because we wanted to observe the kinds of "revisions" readers actually had to make in this prose in order to understand it.

Because regulation prose is so difficult to read, readers must often resort to conscious processing of the prose--we had observed this when the members of our research team first started discussing the meaning of a regulation. Each of us was literally "rewriting" the sentences in order to make sense out of them. Our working hypothesis, then, was that if there were consistent patterns to our readers' "revisions," these might suggest what sort of revisions the writers of these regulations should be making. Protocol analysis was well suited to this task, since it tends to capture just those places where readers encounter difficulty and have to resort to conscious processing of the

prose (Waern, 1979).

Subjects were asked to read and interpret the text, simply trying to say aloud to a tape recorder everything that was going through their minds as they read. In the first study our subjects were given sections of the regulation governing the Small Business and Capital Ownership Development Program (13 CFR Part 124 Section 121.1-121.3) under the jurisdiction of the Small Business Administration (SBA). We selected this regulation because it was ostensibly designed to aid the small-business person, the very person who probably couldn't afford outside legal interpretation of the regulations. We conducted a detailed analysis of the reading protocols of three representative readers: one was a successful entrepreneur with a graduate background in business, the second was a successful businesswoman who had had two SBA loans in the past, and the third was an Assistant District Director for the SBA who had worked in the management assistance office.

B. Coding System

The following major categories emerged from the reader statements.

1. Meta statements. This broad category included the plethora of comments which didn't pertain to the content or meaning of the regulation, such as "This is going to be hard to

read," "All right, I've seen this stuff before," and various other statements about the government, about regulations in general, and about the reader's own process.

2. Content Statements. This body of statements, on which our analysis focused, covered all content-related statements. We divided them into two major sub-categories: Comments and Translations. The category Comments was the more general and included any comment or inference as long as it in some way interpreted the meaning of the regulation (e.g., "All right. They want to keep all this money separate..."). But if a Content statement was tied to some clearly discernable portion of the text, it was coded as a Translation. Translations were, in effect, direct attempts to translate the meaning of some phrase, sentence, or group of sentences into another more understandable form. Comments and Translations are complementary reading statements: one expresses global, general meanings or interpretations, the other operates more locally. For our analysis this distinction was interesting but not crucial. What stood out in our analysis were the three distinctive kinds of Content statements, which we labeled Structural, Retrieval, and Scenario statements.

- Structural statements. This category included any statement concerned with the structure of information in the regulation, whether the reader was using headings to actively search for missing information or using the structure to understand the document, as in "Now see then we go back to a second main point here."

- Retrieval statements. In trying to understand a regulation our subjects often drew on their prior knowledge, including knowledge about the SBA itself or about business and economics, as in the comment, "Oh, they're trying to prevent chaining." Here, the notion of "chaining," which is familiar to people in business but not mentioned in the regulation, let our reader stand back and grasp the unstated, sometimes hidden meaning behind a specific part of a regulation. This category, then, included any statement that indicated explicit retrieval of such prior knowledge from long term memory.

- Scenario statements. The final distinctive kind of statement we saw was also the one which proved to be most interesting. This category included statements in which the reader turns the meaning of a passage into a condition/action sequence or creates a dramatized scenario in which somebody does something. For example, this is a reader's rewrite of the meaning of "ineligible concerns": "Say that if a fellow has a bar and he's selling moonshine which is not taxed,...." Since regulations so rarely express their meaning in this way, these Scenario statements were particulary conspicuous as "revisions" which changed the text from one form of expression to another.

In coding the protocols, all Content statements were classified as either a Comment or Translation. However, a statement could then receive multiple coding if it contained Structural, Retrieval, and/or Scenario features. Our unit of measurement was the clause, with two exceptions. Speakers use frequent relative clauses where writers would merely use a noun, and we wanted to limit our units to clauses that were adding new, independent information. Therefore, restrictive clauses and clauses which were in the subject or object position of a sentence were simply coded as part of the main clause.

C. Results

Figure 1 shows the results of our coding of the three robust categories described above.

Subject	Total Clauses	Total CONTENT Clauses (Comment/Translation ratio)	Structural	Retrieval	Scenario
1	208	127 (49/78)	15	22	75
2	200	114 (70/44)	50	41	42
3	155	102 (28/74)	17	26	65

Figure 1: Categories for Readers' Statements

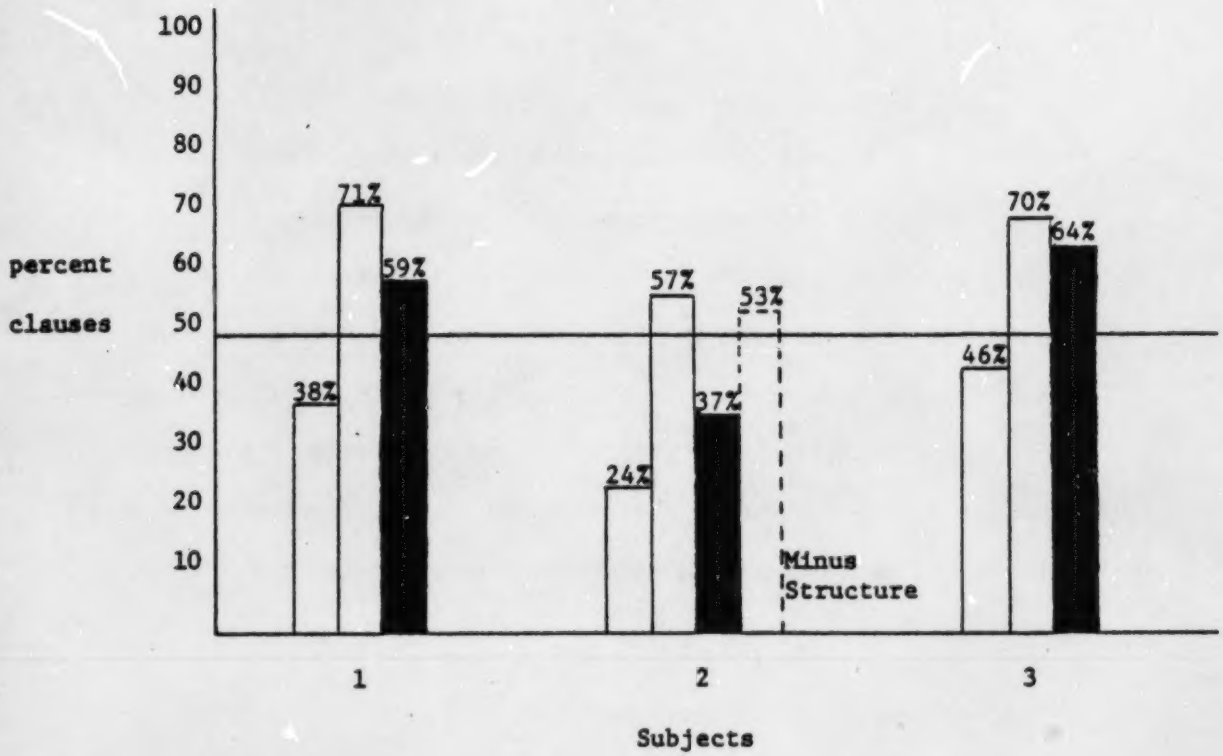
1. The frequency of "Structural" is suggestive. Clearly readers do use structure merely to read the document. The great increase in Structural comments from Subject 2 was due to her attempt to discover whether she was herself eligible for the program. Her concern with structure grew in proportion to her need to actively read the document and search for relevant information. (She was, by the way, never able to find the answer to her question.)

2. All three of our readers showed significant attempts to "Retrieve" and use prior knowledge to understand the prose at hand. As one might expect, the subject with graduate training read pieces of the regulation in terms of economic and business concepts such as "pyramiding." Without that knowledge, parts of the regulation designed to close loopholes (but not explain them, of course) are very difficult to understand. As one might expect, the businesswoman often used her own past experience to interpret parts of the regulation, and interestingly enough, so did the SBA official, as in, "OK, we've made loans to technical schools and so forth, so they--they are eligible."

   Our coding of the explicit use of past knowledge probably reflects only a small portion of what is actually used. It would be naive to assume that we can write clearly about every possible topic to every audience. In order to understand parts of some regulations you may simply have to know a lot about regulations or the situations they are designed to regulate. In rewriting such documents it may make sense to set priorities, determining what level of information must be made clear to the average reader who would use the document and recognizing which levels are dependent on technical or legal knowledge and its language.

3. The most dramatic result from this analysis was the frequency with which our readers resorted to Scenarios in their "revisions" of the regulations, that is, the frequency with which they expressed the meaning of a line or section in the form of a concrete story or event, e.g., with a Condition/Action sequence or with Agents and Actions. As Figure 2 shows, in two of the subjects Scenarios account for over half of all their Content statements (59% and 64%). The proportion for Subject 2 is smaller (37%), due in part to her attempt to search for eligibility information and the resulting high number of comments on structure. If we

eliminate the extra bulk of Structural comments from her total, her proportion of Scenarios becomes, like the others, roughly half of her total comments (53%).



Bars represent: Comment Only / Translate Only / Total Content

Figure 2: Percentage of Scenarios in Content Statements

The Role of Scenarios

Because scenarios account for such a large proportion of our readers' "revisions" (and, as we will see, for such a small proportion of the original prose of the regulation), they deserve more exploration. According to our predictions from the readability principle, we might have expected our readers to break the prose into shorter sentences and to substitute simple words for hard ones: to rephrase the original sentences in simpler ways. Instead, we find them making these much more radical transformations in which a definition is turned into an action and embedded in a context. Below are some examples which compare three versions of the prose our readers produced: a reading of the original, a rephrasing, and a Scenario.

1. Original: No part of such financing shall be used...

   Rephrase: Money can't be used...

   Scenario: If a borrower has that money on hand,...we won't make her a loan so that she can use that money on hand to purchase land or things like that.

2. Original: (Discussion of eligible concerns) Printing: A firm solely engaged in commercial or job printing, if there is no common ownership with a concern ineligible under this paragraph (d)(4) of this section and the printer had no direct financial interest in the commercial success of the material so produced.

   Rephrase: In other words they [this eligible concern] are strictly a job shop printer.

   Scenario: You have something that you need printed, you bring it in, they print it up.

As you can see from Figure 2, the majority of these Scenarios occur in the Translation statements and not, as one might expect, in the more global Comment statements. Readers are using these Agent/Action and Condition/Action structures to restate individual clauses and sentences, not just to comment on or develop the meaning as a whole.

Figure 3 shows the large role Scenarios play in these Translation statements. Based on the predictions of readability, we coded each reader's attempt to clarify a Word, to Rephrase a sentence, or to use a Scenario. (In a few cases noted in the third column, the readers embedded a Scenario in their attempts to rephrase or substitute words.) To our surprise, we found our readers spending relatively little time worrying over "hard words" or trying to parse and restructure big sentences into little ones. Although our readers did use short, simple words and achieved a "standard" level of difficulty on Flesch's Reading Ease scale (i.e., scores of 57, 65, 69), their sentences were long (around 25 words per sentence). But, because these long sentences took the Condition/Action or Agent/Action form of the scenario, they were also quite easy to understand. On Flesch's Human Interest scale they only achieved scores of 20, 29, and 30 out of 100 points (mildly interesting to interesting). In other words, our readers' performance did not violate the cannons of readability, but the protocols suggest that there are better ways to account for the nature of their improvements.

TRANSLATE

Subject	Word Only	Rephrase Only	Scenario (+ Word or Rephrase)	Scenario Only	Scenario Total
1	4 (5%)	18 (23%)	5	51	56 (72%)
2	11 (25%)	8 (18%)	8	17	25 (57%)
3	3 (4%)	19 (26%)	5	47	52 (70%)

Figure 3: Frequency of Scenarios in Translate Statements

What conclusions could we draw from this exploratory study of readers' "revisions"? Readability manipulations would certainly have improved the prose our subjects were reading, but their revisions suggest that they are looking for something more and something different in kind. As the final column in Figure 3 shows, Scenarios account for a large chunk of our readers'

behavior (72%, 57%, 70%). Scenarios represent a more radical kind of transformation than "readability" revisions, because they actually restructure the information. By creating conditions, agents, and actions, they embed the meaning of a regulation in situations and contexts--they make meaning concrete enough to be functional for the reader.

The Nature of Writers' Revisions

Our study of reading protocols revealed a revising principle used by readers. The next logical question to ask is: does this principle apply to written revisions? Do expert writers and revisers also use this principle? Patricia Wright has argued for taking just this sort of approach to solving applied problems in psychology, suggesting that instead of guessing at which variables to test, researchers should start with expert solutions and seek to understand the principles that underlie those successful performances (Wright, 1978). Therefore we used our Scenario hypothesis derived from protocol analysis to compare two regulations: a regulation written in the OLD manner which we knew to be difficult to read (the SBA regulation), and a REVISED regulation which had been generally praised as easy to read (the HEAL or Health Education Assistance Loan regulation). The question we wanted to answer was: do published revisions made by expert regulation writers reflect the Agent/Action focus we saw in the readers' revisions?

Procedure: We compared the OLD and REVISED regulations by selecting comparable segments and counting the number of lines containing clauses with (1) a Human Agent and an Action, or with (2) a human subject if the verb was a state of being verb. In essence, we were looking not simply for active verbs, but human-centered discussions, as in, "The borrower must repay..." and "The Commissioner agrees to pay..." or "A student must be: a citizen, national, or permanent resident in the United States." The alternative was clauses such as, "Eligibility is described...," "Determinations will be made...," or "Membership alone in any group is not conclusive that an individual is socially disadvantaged." We ran this test on two different parts of each regulation--first, on the initial 360 lines which described the program and eligibility and secondly, on two smaller, more specific 160 line sections which told people how to apply for money. As it turned out, each regulation was consistent across its sections, differing by no more than two percentage points. However, the OLD and the REVISED regulations differed considerably, as Figure 4 shows. The SBA writers used Agents and Actions or a human-centered focus in only 33% of the lines, whereas the HEAL writers focused on people and their actions 59% of the time.

Our conclusions, though simple, have broad implications for revision. Our readers interpret prose by organizing it around Human Agents and their Actions--even when those actions are mental ones such as agreeing or deciding. Good government

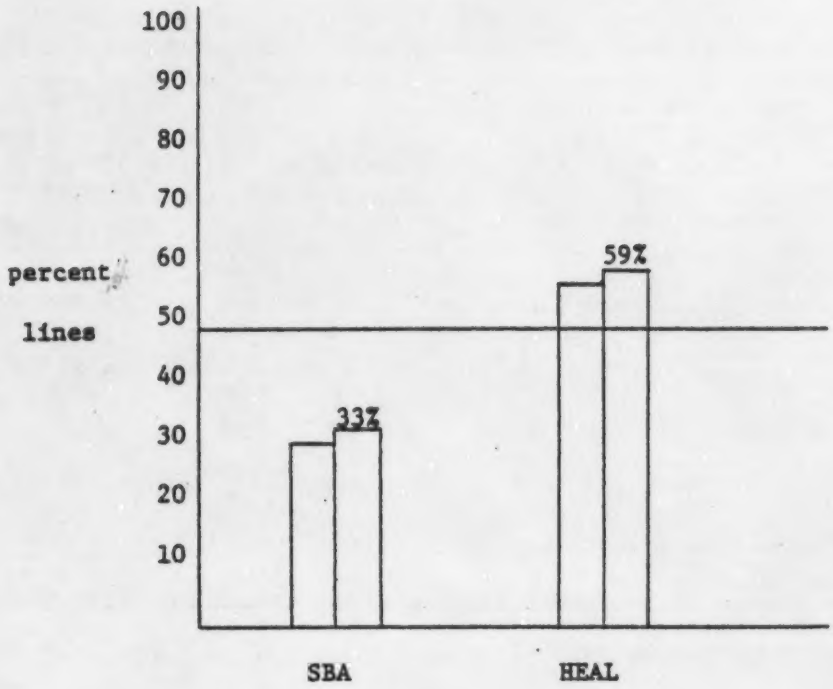


Figure 4: Percentage of Lines with Human Agent/Action Focus

writers/revisers provide that human-centered organization in their prose.

Moving Toward a Revision Principle

On the basis of these exploratory but suggestive studies, it

is possible to step back and characterize some of the key differences between the OLD, hard to read documents studied and the REVISED ones. In speaking of the REVISED versions we will be referring to the revisions made by both the government writer/revisers and our reader/revisers. As before, our goal is to look for a relatively simple but powerful principle for making reader-based revisions.

We have categorized differences between the OLD and REVISED documents in terms of the Writers' Goals and Methods or Writing Strategies (see Figure 5). The primary goal of the OLD regulation seems to be to teach a specified body of information. Like traditional expository writing of the kind found in textbooks, it is rigorously organized around the topic at hand. The REVISED versions, on the other hand, appear designed to help a reader act. This distinction is important since few people ever have occasion to "learn" a regulation or pass a test on it. They read these documents in order to make decisions ("Should I try for a loan?") and to act ("Now how do I apply?"). It would, for example, be ludicrous to study the effectiveness of income tax packets by seeing how much people remember a week after using them. And yet many of the regulations we have seen seem designed to teach information rather than to aid action.

We can clarify these differences further if we look at the rhetorical strategy behind the prose. The OLD version is designed to create a network of information centered around

	GOAL	METHOD OR WRITING STRATEGY
OLD DOCUMENTS	To Teach Information	Develop a Concept-Centered Network — e.g., focus on "Eligibility" → a Definition
REVISED DOCUMENTS	To Help the Reader Act	Develop a Human-Centered Network — e.g., focus on "the Borrower" → a Scenario

Figure 5: Critical Differences Between OLD and REVISED Documents

concepts and key terms. For example, the OLD version has a large chunk organized around the concept of ELIGIBILTY, which is of course a topic of great importance to the agency. The REVISED regulations, however, try to create a human-centered network of information focused on Human Agents and their Actions, such as borrowers who live in the United States and promise to repay

loans. A useful analogy here, although it is only an analogy, is the distinction often made between semantic and episodic memory. While semantic memory is organized around our knowledge of words and their associated meanings, episodic memory is organized around the events and actions of our own lives. The issue here is not memory, but the way one chooses to organize a body of information in the first place.

We can characterize the difference between these two writing strategies by saying that the OLD regulations attempt to build a definitional network held together by key words and concepts (that is, ones which are key to the agency who administers the regulation). Its prose style derives from patterns of classical definition with particular use of categorical definition and the listing of attributes. Nominalizations predominate along with passive constructions and non-human subjects of clauses. The style and writing strategy we see in the REVISED versions are quite different. Information is organized around Agents and Actions placed in a Context: it is this strategy which we have been calling the Scenario principle.

Although such an organization seems to be the height of common sense, it is not fully or commonly recognized in most revision methods. For example, a Scenario structure is a more specific and rigorous constraint than the "human interest" factor scored by Flesch's readability formula which counts the number of persons, personal pronouns, and gender specific references.

Scenarios are concerned with the active role humans play in the world that the regulation depicts. Furthermore, the Scenario principle goes beyond the recommendations made by psycholinguists and technical writing texts which urge writers to transform passive expressions. The research on passive generally pays little attention to whether the subject of the new active sentence is a person or a thing. However, the Scenario principle says that in documents such as regulations, or contracts which demand performance, expert revisers create a Human Agent.

To sum up, we think that this exploratory analysis of the role of Agents/Actions and Context points us toward a principle for revising documents such as regulations. The Scenario principle meets our first criterion of being simple and powerful--this change alone can make a significant difference in the prose. Secondly, unlike readability measures, it can deal with overall organization and it can produce reader-based revisions. This is not, of course, to discount the importance of revising for familiar words and shorter sentences, but to suggest a companion principle that does what readability formulas can't do. According to the Scenario principle as we see it, readers and expert writers are attempting to do two things:

1. They are attempting to build Scenarios, that is, to organize information around Human Agents and their Actions;

2. They are attempting, to varying degrees, to "instantiate" those Scenarios, that is, to provide the context and concrete instances that make a general

rule more understandable.

Practical Revision Strategies

A principle for revising functional documents is only as good as it is useful. Although the notion of a Scenario principle obviously needs further and more detailed linguistic analysis, the overriding purpose of this study was to look for a practical principle by which people could make reader-based revisions. Such a principle needs to be flexible (unlike the rule, "Always avoid the passive voice" which is not always appropriate). And it needs to be powerful, that is, able to draw on the intuitions of practicing writers in business and government and to lead to revision of all levels, i.e., not only phrases and sentences, but the overall organization of the document. In the rest of this paper, then, we would like to follow out some of the practical implications of the Scenario principle and look at revision techniques under two headings:

- Structuring information around the reader;
- Instantiating the Scenario by providing a context for meaning.

A. Structuring Information Around the Reader

Structuring prose with the reader in mind can be done in many ways on many levels. We consider three levels: the top-level structure, the local structure, and the grammatical

structure.

1. The Top-Level Structure (major ideas units, headings, overall organization)

- The top-level structure should accommodate the search process of readers searching for needed information or using the regulation as a reference document. Currently many headings are meaningful to an agency reader, but not to a user reader. Or, they contain "miscues" which actually mislead (Swarts, Flower, & Hayes, 1980).

- Because such functional documents allow people to make decisions and take actions, they should be organized where possible around the actions people take rather than around key terms. For example, a discussion organized around the decision, "How to decide if you are eligible," will be more understandable than a definitional discussion of eligibility. The information would remain the same, but its comprehensibility would increase.

- The top-level structure should in many cases be designed to answer the reader's questions. Good writing creates a context for its information, introducing new facts in the light of old, answering questions, and fulfilling expectations. In general, with functional documents this is particularly important since readers normally come to them with a purpose; they read to find answers to their own questions. We could imagine that there is a general set of questions people bring to most documents, such as: What is this about? To whom does it apply? What do I have to do? However, elementary research into the needs of actual users can often reveal many questions which are specific to a particular document. For example, Siegel & Gale, a firm which specializes in writing and designing commercial documents, conducts such research prior to revision. On one loan agreement, they discovered that one of the questions borrowers most frequently asked was, "Can I repay the loan early?" Yet this information was buried in the previous loan form because it was not critical to the writer, the bank.

A Test Case. We wished to test the notion that good writers

use the Scenario principle even in the headings of documents. Therefore we ran a small test to see if judges, using the Scenario principle alone, could successfully discriminate the headings of OLD and REVISED documents from a scrambled list of headings. The headings came from four documents: an OLD (the SBA) and REVISED (HEAL) regulation and an OLD and REVISED version of a Personal Liability Catastrophy Policy. The judges' instructions were: mark as OLD any heading that appears to be trying to define a term that is of importance to the agency or company; mark as REVISED any heading which (1) contains an Agent or Action or (2) is an answer to a question the reader would have. Examples of OLD headings included: Personal Liability; Occurrence; Social Disadvantage. REVISED ones included: Who is an Eligible Student Borrower? The Loan Application Process; Can This Policy Be Cancelled?

Out of a total of 63 headings, our four subjects were able to identify correctly 20 of the 27 headings from the REVISED versions. This high 74% correct suggests that the Scenario principle can account for many of the changes good writers make. Why wasn't it 100%? One possibility is that even revised regulations and policies aren't as good as they might be.

2. On a more local level the Scenario principle suggests other techniques for structural revision.

- When appropriate, writers should use condition/action statements or old/new patterns. Here is an example

from the HEAL regulation which defines the "borrower's obligations" in terms of a concrete situation and the subsequent actions people take: "If the lender agrees to make the loan and if the Commissioner of Education approves the loan for HEAL insurance, the applicant must then sign a promissory note."

- When possible, writers should use operational definitions which describe the meaning of a concept in practical, operational terms. Such definitions, much like a recipe, spell out the procedures for creating or attaining the thing defined. For example, HEAL defines eligibility in such a way that the reader can clearly determine whether he or she is eligible, and if not eligible, see the steps one would have to take to become so: "Eligibility: He or she must be carrying or plan to carry, during the period for which the loan is intended, the normal workload of a full-time student as determined by the school."

3. Finally, the Scenario principle also offers suggestions at the level of grammatical structure.

- Writers should try to use agents and actions in the subject and verb positions of the sentence. In an important article on the revision of functional documents in general, Williams has argued for the value of making agents and grammatical subjects coincide (Williams, 1979).

- In regulations, at least, we feel it is particularly important to make that agent a Human Agent. Generally such documents are designed to regulate the actions of the reader, but it is difficult for the reader to identify his or her own role as agent when the prose doesn't. For example, one section was organized around a long list of failures (such as "failure to reasonably pursue commercial and competitive business in accordance with the business plan projection"), when the real purpose was to threaten dire consequences "if you [the reader] fail to pursue commercial, competitive business according to the business plan you projected." On this count the HEAL regulation set a fine precedent by setting up in the prose a Commissioner who "approves" loans, "signs" documents, and "must be satisfied" at various times. The SBA's constant reference to "an applicant concern" and "a section 8(a) business concern" was a step in the right direction,

but one which many of our readers found confusing and only on the borderline of "human."

B. Instantiating the Scenario

A second major way in which writers can act on the Scenario principle is to try to instantiate the Scenario itself, that is, to give the reader a sense of the instances which underlie the principle. We have chosen the term "instantiate" to indicate a process that is broader than simply giving an example. Like the terms "hypostatize" or "reify" in philosophy, it suggests the act of filling in the slots in an abstract framework in much the way people use familiar mental scripts or schemas to make sense of new information. Such schemas are only empty structures that become meaningful when they are "instantiated" with a specific, concrete reality.

1. A simple but space-consuming way to instantiate a scenario is to provide extended examples or cases. In revising the insurance policy noted above, Siegel & Gale drew the following case from the company's files to vividly illustrate the notion of "Broader Coverage" offered by the policy:

> You've boarded your neighbors' poodle while they're away on vacation. You're careless and the poodle runs away and gets lost. Your neighbors insist on your paying for the loss. If he was an ordinary poodle worth say $400, you pay $200 and we pay $200. But if he was a prize-winning show dog worth $4,000, you pay $250 and we pay $3,750.

Another, briefer version of this would be a short example interpolated into the prose, as in, "In addition, certain non-students (such as doctors serving as interns or residents)...."

There are, however, problems in using extended examples since some research has suggested that given the opportunity, readers prefer the example to the discussion and mistakenly assume that the example represents the range of the given term's meaning. In their income tax forms, Sweden, however, has dealt with this problem by using elaborate scenarios and examples which cover the major typical categories of people who fill out income tax forms. Due in part to the clearly dramatized, Sunday supplement style of the instructions which accompany the tax forms, a large percentage of Swedish taxpayers are able to fill out their own forms. The minority who don't fit in the typical cases must then turn to more complex instructions (Anna-Lena Ericsson, personal communication, 1979).

2. Where possible, writers can use detail and specifics to define the meaning of a term. For example, instead of saying that it offers "insurance of education loans for certain health professions," the HEAL regulation says, "The health education assistance loan (HEAL) program is a program of Federal insurance of educational loans designed for students in the fields of medicine, osteopathic medicine, dentistry, veterinary medicine, optometry, podiatry, pharmacy, and public health." Although such

a detailed sentence might fail a readability test because of length, most people would not find it difficult to read. In this case, the length is what makes it comprehensible. The shorter, alternative version we offered above would be easy to process as a piece of prose, but it would not answer questions a reader might have. Nor is it clear that a person reading it would really understand the meaning of that statement without the instantiation--the examples and particulars which give body to the generalization. In this case we mean "understand" in a pragmatic way: the reader might know the meaning of the words "health professions," but not be able to generate the particulars they stand for or recognize that they apply to himself.

Some Unanswered Questions

This exploratory study has highlighted the importance of context and in doing so raises a difficult and perplexing question about the nature of context and instantiation: how much is enough?

Here is an example of how instantiation normally works in mathematics. A general principle or rule without any instantiation would be R x T = D or Rate x Time = Distance. We could instantiate it by saying 90 mph times 2 hours equals 180 miles, or provide an even more vivid context by saying, "If you head out of New York going 90 mph, you'll be in Boston (220 miles away) in less than 2-1/2 hours." In logic one would say that the

first version, R x T = D, represents an "intentional" definition. Like a biological definition of mammal, it lists the rules or attributes which define the term. An "extentional" definition, by contrast, lists examples, such as dogs, cats, and platypuses. Clearly, people need both aspects of a definition, the principle and the instance. The question is, how much instantiation is enough? For readers of public documents it may be much more than we think.

This study raises another provocative question, which is: how far should we generalize the value of the Scenario principle? Our working hypothesis in this study was that Scenarios were particularly appropriate for functional documents, such as regulations, in which people needed to read in order to act. As one way of testing this hypothesis we collected more reading protocols from 20 subjects who read a passage of difficult prose under two conditions. In the first condition the readers were told they were to verbalize their thoughts as they tried to read and understand a passage from a textbook. In the second condition the subjects were again instructed to verbalize their thoughts, but were told that the passage (same passage) was one which they would have to explain to someone else. One hypothesis in this experiment was that the readers with the prospect of explaining these rules would behave like regulation readers and resort to more Scenarios than the textbook readers. However, in fact, both groups behaved much like the first group of regulation

readers, using Scenarios in 40% to 50% of their clauses. Perhaps Scenarios are simply a way people translate any sort of difficult prose in an attempt to make it comprehensible. Support for this notion and for the general importance of Scenarios comes from an unlikely source. According to many rhetoric and composition handbooks (Hairston, 1974; McCrimmon, 1963), Scenarios are a pervasive feature of English prose, although from the handbook point of view they are a fault and are labeled as a "faulty complement" or "faulty predicate." An example would be, "Subversive propaganda is when someone tries to turn people against their country." Propaganda cannot technically be a "when," but perhaps the persistence of this fault of style is simply an indication of the importance of Scenarios as a pattern with which people think.

Conclusion

The clear language movement with its demand for understandable public documents has become a public issue. If clear writing is going to be produced, regardless of legislation, revisers need practical principles which are more powerful than readability alone. Such principles, we feel, should improve the comprehensibility, or what Wright has called the usability of the document (Wright, 1979). This means they must be able to diagnose and solve problems which involve the structure of the information. In addition, any such principle must seem

self-evident and intuitively right to the people who will be using it; it must be easy to use and to teach or it would simply be another piece of "academic" unused wisdom. The Scenario principle is one promising candidate which, as a complement to readability, could help us write more reader-based public documents.

REFERENCES

Ålderman, G. H. Journal of educational research, 1926, 13, 11.

Bobrow, D. G., & Norman, D. A. Some principles of memory schemata. In D. G. Bobrow & A. Collins (Eds.), Representation and understanding: Studies in cognitive science. New York: Academic Press, 1975.

Bobrow, R. J., & Brown, J. S. Systematic understanding: Synthesis, analysis, and contingent knowledge in specialized understanding systems. In D. G. Bobrow & A. Collins (Eds.), Representation and understanding: Studies in cognitive science. New York: Academic Press, 1975.

Charrow, V. R. What is "plain English," anyway? (Document Design Center Report). Washington, D.C.: American Institutes for Research, December 1979.

Clark, H. H., & Clark, E. V. Psychology and language: An introduction to psycholinguistics. New York: Harcourt Brace Jovanovich, Inc., 1977.

Flower, L. Writer-based prose: A cognitive basis for problems in writing. College English, September 1979, 41, 19-37.

Flower, L. S., & Hayes, J. R. A process model of composition (Document Design Project Technical Report No. 1). Pittsburgh, Pa.: Carnegie-Mellon University, August 15, 1979. Also in L. W. Gregg & E. R. Steinberg (Eds.), Cognitive processes in writing. Hillsdale, N.J.: Lawrence Erlbaum, in press.

Hairston, M. A contemporary rhetoric. Boston: Houghton Mifflin, 1974.

Kern, R. P., Sticht, T. G., Welty, D., & Hauke, R. N. Guidebook for the development of army training literature. Washington, D.C.: U. S. Army Institute for the Behavioral and Social Sciences, Human Resources Research Organization, November 1976.

Klare, G. R. The measurement of readability. Ames, Ia: Iowa State University Press, 1963.

Kniffen, J. D. The practical application of readability/ comprehensibility research to the production of more usable military technical materials. Hunt Valley, Md.: Westinghouse

Electric Corporation, 1978.

McCrimmon, J. M. Writing with a purpose. Boston: Houghton Mifflin Company, 1963.

Redish, J. C. How to draft more understandable legal documents. In Drafting documents in plain language (Course Handbook No. A4-3034). New York: Practising Law Institute, 1979.

Redish, J. C. Readability. In D. B. Felker (Ed.), Document design: A review of the relevant research. Washington, D.C.: Document Design Project, American Institutes for Research, 1980.

Redish, J. C., & Racette, K. Teaching college students how to write: Training opportunities for document designers (Document Design Project Report). Washington, D.C.: American Institutes for Research, November 1979.

Steinberg, E. R. A new era for research in composing. Unpublished manuscript, Carnegie-Mellon University, 1979.

Swarts, H., Flower, L. S., & Hayes, J. R. How to misread a Federal regulation. In D. B. Felker (Chair), Some research on document design and document designers. Session presented at the meeting of the American Educational Research Association, Boston, 1980.

Waern, Y. Thinking aloud during reading: A descriptive model and its application (Department of Psychology Report No. 546). Stockholm, Sweden: University of Stockholm, April 1979.

Williams, J. M. Defining complexity. College English, February 1979, 40, 595-609.

Wright, P. Reading to learn. Chemistry in Britain, 1968, 4, 445-450.

Wright, P. Feeding the information eaters: Suggestions for integrating pure and applied research on language comprension. Instructional Science, 1978, 7, 249-312.

Wright, P. Usability: The criterion for designing written information. Paper presented at the NATO Conference on Visible Language.2, Toronto, September 1979.